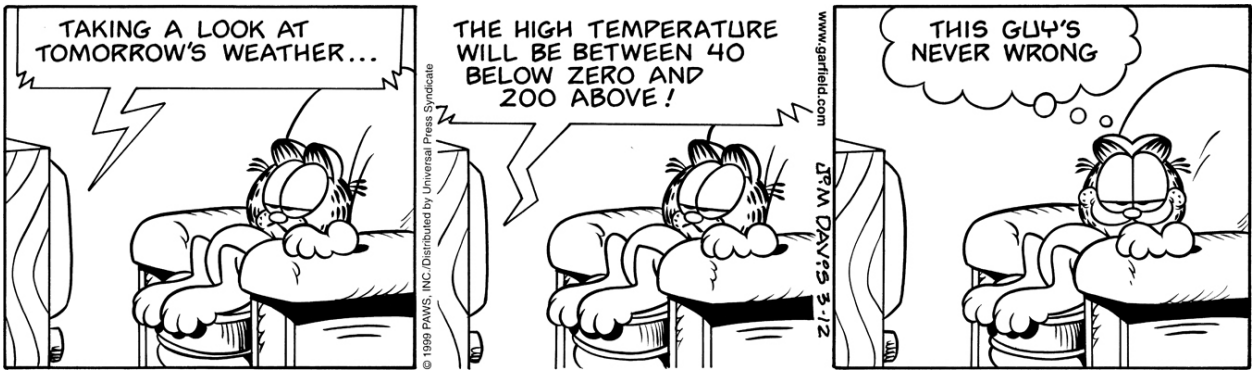
Chapter 8: Estimating with Confidence

Key Vocabulary:

- point estimator
- point estimate
- confidence interval
- margin of error
- interval
- confidence level
- random
- normal
- independent
- four step process
- level C confidence interval
- degrees of freedom
- standard error
- one -sample z interval
- t distribution
- t-procedures
- one-sample t interval
- robust



8.1 Confidence Intervals: The Basics (pp.615-643)

1. A *point estimator* is a statistic that…

2. The value of the point estimator statistic is called a _____ and it is our "best guess" at the value of the _____.

3. Summarize the facts about *sampling distributions* learned in chapter 7:

   Shape

   Center

   Spread

4. In statistics, what is meant by a 95% *confidence interval*?

5. A confidence interval takes the form of : "**estimate ± margin of error**"
   where: estimate =
   margin of error =

6. Define a *level C confidence interval.*

7. What information does the margin of error provide?

8. Sketch and label a *95% confidence interval* for the standard normal curve.

9. In a sampling distribution of $\bar{x}$, why is the interval of numbers between $\bar{x} \pm 2s$ called a *95% confidence interval*?

10. Sketch and label a *90% confidence interval* for the standard normal curve.

11. *Interpret a Confidence level*: "To say that we are 95% confident is shorthand for …..

12. Explain how to interpret a *Confidence interval*.

13. Does the confidence level tell us the chance that a particular confidence interval captures the population parameter? If not, what does it tell us?

14. What does the *critical value* depend on?

15. Write the *form* for calculating a confidence interval as shown on page 478.

16. Why do we want high confidence and a small margin of error?

17. Explain the two conditions when the margin of error gets smaller.

18. State the three **conditions for constructing a confidence interval** for p or μ.

- Random

- Normal

- Independent

19. What are the two important reminders for constructing and interpreting confidence intervals?

8.2 Estimating a Population Proportion (pp.484-494)

1. In statistics, what is meant by a *sample proportion:* $\hat{p}$?

2. Give the mean and standard deviation for the sampling distribution of $\hat{p}$?

3. How does the standard deviation differ to to standard error for the sampling distribution of $\hat{p}$?

4. Describe the sampling distribution of a sample proportion $\hat{p}$ as learned in section 7.2.
   - Shape
   - Center
   - Spread

5. Define *standard error.*

6. In general what is meant by the standard error of a statistic?

7. How do you calculate the standard error of $\hat{p}$?

8. What is the formula for a *one-sample z interval for a population proportion*? Describe how to construct a level C confidence interval for a population proportion.

9. Describe the four step process on how to contruct and interpret a confidence interval.

   - State
   - Plan
   - Do
   - Conclude

10. What formula is used to determine the sample size necessary for a given margin of error?

11. What conditions must be met in order to use *z procedures* for inference about a proportion?

12. What does z^* represent?

13. What is the value of z^* for a *95% confidence interval*? Include a sketch.

14. What is the value of z^* for a *90% confidence interval*? Include a sketch.

15. What is the value of z^* for a *99% confidence interval*? Include a sketch.

8.3 Estimating a Population Mean (pp.499-515)

1. What is the formula for a *one-sample z interval for a population mean*? Describe how to construct a level C confidence interval for a population mean.

2. What is the formula for the margin of error of the confidence interval for the population mean μ?

3. How can you arrange to have both high confidence and a small margin of error?

4. Describe the three steps for choosing a sample size for a desired margin of error when estimating μ.

5. What happens to the *margin of error* as *z** gets smaller? Does this result in a higher or lower confidence level?

6. What happens to the *margin of error,* as σ gets smaller?

7. What happens to the *margin of error,* as *n* gets larger? By how many times must the sample size *n* increase in order to cut the *margin of error* in half?

8. The formula used to determine the sample size *n* that will yield a confidence interval for a population mean with a specified margin of error *m* is $z^* \frac{\sigma}{\sqrt{n}} \le ME$. Solve for *n*.

9. It is the size of the _____ that determines the margin of error. The size of the _____ does not influence the sample size we need.

10. Complete the Check Your Undertanding on page 501.

11. How do you calculate the *degrees of freedom* for a *t distribution*?

12. What happens to the *t distribution* as the *degrees of freedom* increase?

13. How would you construct a t distribution?

14. Describe the differences between a *standard normal distribution* and a *t distribution*.

15. Describe the similarities between a *standard normal distribution* and a *t distribution*.

16. What is the formula for the *standard deviation* of the sampling distribution of the sample mean $\bar{x}$?

17. What is the *standard error* of the sample mean $\bar{x}$?

19. Describe how to construct the *one-sample t interval for a population mean*?

20. Summarize the *three conditions for inference about a population mean:*
    - Random
    - Normal
    - Independent

21. Inferences for *proportions* use _____ and inferences for *means* use _____.

22. What does it mean if an inference procedure is **robust**?

23. If the size of the SRS is less than 15, when can we use *t procedures* on the data?

24. If the size of the SRS is at least 15, when can we use *t procedures* on the data?

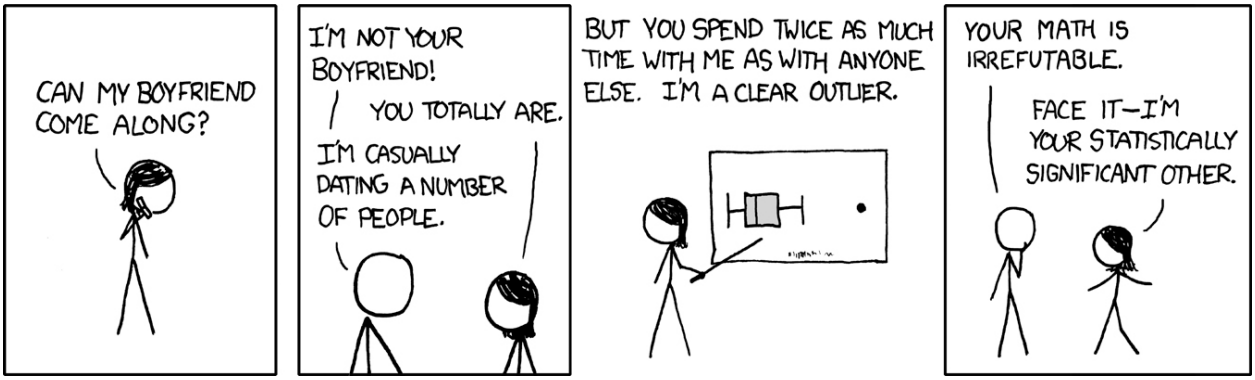
25. If the size of the SRS is at least 30, when can we use *t procedures* on the data?

26. Summarize the details of the four step procedure for estimating *p*:
    - State
    - Plan
    - Do
    - Conclude

Chapter 9: Testing a Claim

Key Vocabulary:

- Significance test
- Null Hypothesis
- Alternative Hypotheses
- One sided alternative
- Two sided alternative
- p-value
- α level
- significance level
- one-sample z test
- test statistic
- one sample t test
- paired data
- four-step process
- statistically significant
- Type I Error
- Type II Error
- Power
- Degrees of freedom
- t-distribution
- paired t procedures



9.1 Significance Tests: The Basics (pp.528-543)

1. What is a *significance test*?

2. What is the difference between a *null and an alternative hypothesis*? What *notation* is used for each?

3. Explain the differences between one-sided and two-sided hypotheses. How can you decide which one to use?

4. What form does the null and alternative hypothesis take in significance testing?

5. Hypotheses always refer to a _____, not to a _____.

6. In statistics, what is meant by the *P-value*? What does a P-value measure?

7. If a *P-value* is small, what do we conclude about the *null hypothesis*?

8. If a *P-value* is large, what do we conclude about the *null hypothesis*?

9. What are *common errors* students make in their conclusions of P-values?

10. On what evidence would we *reject the null hypothesis*?

11. On what evidence would we *accept the null hypothesis* (ie. fail to reject the null hypothesis)?

12. What is meant by a *significance level*?

13. Explain what it means to say that data are *statistically significant*.

14. How small should the *P-value* be in order to claim that a result is *statistically significant*?

15. When using a *fixed significance level* to draw a conclusion in a statistical test what can be concluded when the P value is $< \alpha$ and $\geq \alpha$?

16. What two circumstances guide us in choosing a level of significance?
    -
    -

17. What is a *Type I Error*?

18. What is a *Type II Error* ?

Decision based on sample	Truth about the population: H_0 true	Truth about the population: H_a true
Reject H_0	Type I error	Correct decision
Fail to reject H_0	Correct decision	Type II error

19. Which error is worse, Type I or Type II?

20. Complete the *Check Your Understanding on page 539*.

21. What is the relationship between the *significance level* α and the probability of *Type I Error*?

22. How can we reduce the *probability* of a Type I error?

23. What is meant by the *power* of a significance test?

24. What is the relationship between *Power and Type II Error*? Will you be expected to calculate the power on the AP exam?

25. What *four factors* affect the power of a test? Why does this matter?

26. Describe the three influences that must be verified before deciding on how many observations are needed in a study.

    - Significance Level
    - Practical Importance
    - Power

9.2 Tests about a Population Proportion (pp.549-561)

1. Summarize the three conditions that must be checked before carrying out significance tests:
   -
   -
   -

2. State the general form of the "*test statistic*".

3. What does the *test statistic* measure? Is this formula on the AP exam formula sheet?

4. Describe the four step process for signifigance tests. Explain what is required at each step.
   - State
   - Plan
   - Do
   - Conclude

5. What test statistic is used when testing for a *population proportion*? Is this on the formula sheet?

6. Summarize the one-sample z test for a proportion and sketch the three possible H_a's.
   - Choose…
   - To test…
   - Find…
   - Use this test…
   - If Normaility is not met

7. What happens when the data does not support H_a?

8. If asked to carry out a signifigance test and there is no α provided, what is recommended?

9. Can you use confidence intervals to decide between two hypotheses? What is the advantage to using confidence intervals for this purpose?

10. Why don't we always use confidence intervals?

9.3 Tests about a Population Mean (pp.565-585)

1. What are the three conditions for conducting a *significance test for a population mean*?

2. What *test statistic* do we use when testing a *population mean*? Is this formula on the AP exam formula sheet?

3. How do you calculate *p-values using the t-distributions*?

4. What do you do if the *degrees of freedom* you need is not in table b?

5. How do you find *p-values* when carrying out a signifigance test about a population mean on the *calculator*?

6. For a **one-sample** t- test for a **population mean**, state:

- H_0

- the three possible H_a's (with small sketches to illustrate)

- What is the t test statistic and how is it interpreted?

- Under what conditions can this test be used…

7. In terms of rejecting the hypothesis H_0, how is a significance test related to a confidence interval on the same population?

8. Use your calculator to find the p value (tcdf command) for the example Healthy Streams. What is that p-value?

9. When using technology for the *"DO"* part of the four step process, what is recommended on page 573?

10. Work through the *Juicy Pineapple example* on page 574. Use a calculator to find the exact P-value. Why is *tcdf* mulitplied by 2?

11. Why is the difference between using the calculator versus Table b when finding the p-value in this example?

12. Do we have encough evidence to reject H_0 in the Juicy Pineapple example? Explain.

13. Read the Check Your Understanding on page 577 and answer questions 1 and 2.

14. What is *paired data*?

15. What information would lead us to apply a *paired t-test* to a study, and what would be the statistic of interest?

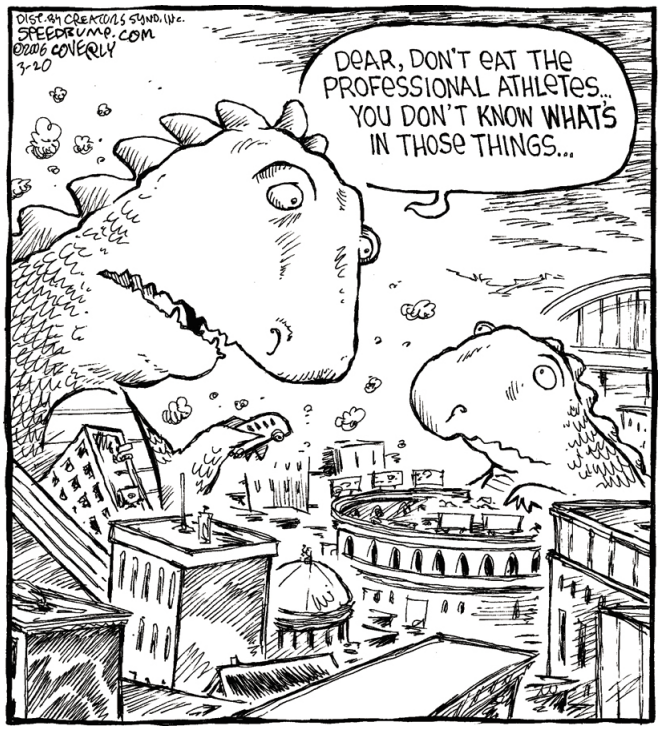
16. In the example, *Is Caffeine Dependence Real*, explain the difference in the "Do" procedures for this example versus the *Juicy Pineapple* example.

17. Describe the four points to be aware of when *interpreting* signifigance tests.

Chapter 10: Comparing Two Populations or Groups

Key Vocabulary:

- difference between two proportions
- two sample z interval for proportions
- two sample z test for difference between two proportions
- two sample z statistic
- two sample t statistic
- pooled combined sample proportion
- standard error
- randomization distribution
- paired t-test
- two sample t test for means
- two sample t interval for means
- difference between two means
- pooled two sample t statistic



10.1 Comparing Two Proportions (pp. 604-618)

1. Summarize the three properties of a sampling distribution of a sample proportion:
   - Shape
   - Center
   - Spread

2. What are the shape, center, and spread of the sampling distribution of $\hat{p}_1 - \hat{p}_2$? Provide the formulas for the mean and standard deviation.
   - Shape
   - Center
   - Spread

3. What conditions need to be met for the sampling distribution of $\hat{p}_1 - \hat{p}_2$?

4. Give the formula for the *standard error* when calculating a confidence interval for $\hat{p}_1 - \hat{p}_2$, and define each variable in the equation.

5. What is the confidence interval for $\hat{p}_1 - \hat{p}_2$?

6. What conditions must be met in order to use the Two-sample z Interval for a Difference between Two Proportions?

   - Random

   - Normal

   - Independent

8. Use the example, *Teens and Adults on Social Networking Sites*, to outline how to construct and interpret a confidence interval for the difference between two proportions, $p_1 - p_2$.

9. State the null hypothesis for a *two proportion significance test.*

10. What does $\hat{p}_c$ represent, and how is it calculated?

11. Why do we *pool* the sample proportions?

12. Give the formula for the *two-proportion z-statistic*, and define each variable in the equation.
13. Is this on the formula sheet? What does the test statistic measure?

14. State and use diagrams to illustrate the three possible alternative hypotheses for a *two proportion z-test.*

15. What are the *conditions* for conducting a two-sample *z* test for a difference between proportions?

16. How are these *different* than the conditions for a one-sample *z* interval for *p*?

17. Describe the *randomization distribution*.

18. What must you be careful about when *defining parameters* in experiments? How can this be avoided?

19. Can you use your calculator for the *Do* step? Are there any drawbacks?

20. What are the calculator commands for the two-sample z test and interval for $\hat{p}_1 - \hat{p}_2$?

10.2 Comparing Two Means (pp.627-648)

1. Summarize the three properties of a sampling distribution of a *sample mean*:
   - Shape
   - Center
   - Spread

2. What are the shape, center, and spread of the sampling distribution of $\bar{x}_1 - \bar{x}_2$? Give the formula for the mean and standard deviation.
- Shape
- Center
- Spread

3. What are the conditions for the sampling distribution of $\bar{x}_1 - \bar{x}_2$?

4. Give the formula for the *two-sample t-statistic*, and define each variable in the equation.

5. Is this on the formula sheet? What does it measure?

6. What is the standard error of $\bar{x}_1 - \bar{x}_2$? Is this on the formula sheet?

7. What distribution does the two-sample *t* statistic have?
8. Why do we use a *t* statistic rather than a *z* statistic?

9. Without using technology, how do you estimate the degrees of freedom when using two-sample t-procedures?

10. How do you calculate the confidence interval for $\mu_1 - \mu_2$?

11. In a *two-sample t interval* problem, what conditions must be met for comparing two means?

12. What are the conditions for conducting a two-sample *t* test for $\mu_1 - \mu_2$?

13. Draw a sketch of the three possible scenarios for the alternative hypothesis.

14. Describe the *Normal Condition* when using the two sample t procedures.

15. What calculator commands are used for a two-sample t test and interval for $\mu_1 - \mu_2$?

16. How do you proceed when using two-sample t procedures to check the Normal Condition in the following cases:
    - Sample size less than 15
    - Sample size at least 15
    - Large samples

17. In a two-sample problem, must/should the two sample sizes be equal?

18. When doing two-sample *t* procedures, should we pool the data to estimate a common standard deviation? Is there any benefit? Are there any risks?

Chapter 11: Inference for Distributions of Categorical Data

Key Vocabulary:

- one way table
- chi-square test for goodness of fit
- chi-square statistic
- expected count
- observed count
- chi square distribution
- degrees of freedom
- chi-square distribution
- components of chi-square
- cell counts
- r x c table
- chi square test for homogeneity
- chi square test for association/ independence

11.1 Chi-Square Goodness of Fit Test (pp.678-690)

1. What is a *one-way table*?

2. What is a *chi-square goodness-of-fit test*?

3. What is the difference between the notation X^2 and χ^2?

4. State the general form for the *null hypotheses* for a χ^2 goodness of fit test.

5. State the general form for the *alternative hypotheses* for a χ^2 goodness of fit test.

6. How do you calculate the *expected counts* for a chi-square goodness-of-fit test? How should you round the answer for the expected counts?

7. What is the shape of a *chi-square distribution*? What happens to the shape as the degrees of freedom increases? (Illustrate with a diagram)

8. Describe the *center and spread* of the chi-square distributions.

9. What is the *chi-square test statistic*? Is it on the formula sheet? What does it measure?

10. How many degrees of freedom does the *chi-square distribution* have?

11. What is the *rule of thumb* for all expected counts in a chi-square goodness of fit test?

12. What conditions must be met in order to use the *goodness of fit test*?

12. How do you calculate *p*-values using chi-square distributions?

14. Can you use your calculator to conduct a chi-square goodness-of-fit test? If yes, what are the calculator commands?

15. What is meant by a *component* of chi-square?

16. What does the *largest component* of chi-square signify?

17. Why is it necessary to perform *follow-up analysis* to a chi-square test?

14.2 Inference for Relationships (pp.696-721)

1. What is the *hypothesis* for a test of homogeneity?

2. Describe the complications with *multiple comparisons*? How are they overcome?

3. Explain how to calculate the expected counts for a test that compares the distribution of a categorical variable in multiple groups or populations.

4. Write the *formula* for the Chi-square test statistic? Is this on the AP Exam formula sheet?

5. What does the Chi-square *test statistic measure*?

6. What information is contained in a *two-way table* for a Chi-square test?

7. How many *degrees of freedom* does a chi-square test for a two-way table with *r* rows and *c* columns have?

8. What *requirements* must be checked before carrying out a Chi-square test for Homogeneity?

9. State the null and alternative hypothesis for the Chi-square test for Homogeneity?

10. Can you use your calculators to do a Chi-square test of homogeneity? If yes, what are the calculator commands?

11. Summarize how to carry out a *Chi-square Test for Homogeneity of Populations*:

12. Explain how and when to conduct a *follow-up analysis* for a test of homogeneity?

13. What does it mean if two variables have an *association*?

14. What does it mean if two variables are *independent*?

15. State the *null and alternative hypotheses* for a Chi-square test for Association/Independence.

16. How is a test of association/independence *different* than a test of homogeneity?

17. How do you calculate *expected counts* for a test of association/independence?

18. Summarize how to carry out a Chi-square Test for Association/Independence:

19. What are the *conditions* for a test of association/independence?

20. When should you use a *chi-square test* and when should you use a *two-sample z test*?

Chapter 12: More about Regression

Key Vocabulary:

- sample regression line
- true regression line
- t interval for slope
- standard error of slope
- t test for slope
- standardized test statistic
- standard error
- exponential model
- power model
- logarithmic model

12.1 Inference for Linear Regression (pp.739-757)

1. What is the difference between a *sample regression line* and *population (true) regression line*?

2. Explain the *sampling distribution* of *b*?

3. Give the equation for the *true regression line*, and state what each component of the equation represents.

4. Summarize the *conditions* for regression inference:

- L
- I
- N
- E
- R

5. Explain how to *check the conditions* for regression inference:

- L
- I
- N
- E
- R

6. Record the formula for the *standard error of the slope*? Define the variables.

7. What is the formula for the *t-interval of the slope* of a least-squares regression line? Is this on the AP exam formula sheet?

8. What is the formula for the *t-test for the slope* of the population regression line? Is this on the AP exam formula sheet?

9. Describe the distribution of the *standardized test statistic* $\frac{b-\beta}{\text{SE}_b}$.

10. What is the formula for constructing a *confidence interval for a slope*?

11. What calculator commands are used to get the value of t^*?

12. Can you use your calculator to conduct a test and confidence interval for the slope?

12.2 Transforming to Achieve Linearity (pp.765-783)

1. What does it mean to *transform data*?

2. What is a *power model*?

3. Give three *example*s of power models?

4. Aside from power transformations, how can you *linearize an association* that follows a power model in the form $y = ax^p$?

5. Describe a *logarithmic model*. Give two examples.

6. Describe an *exponential model*. Give two examples.

7. Describe the two methods used to linearize a relationship that follows an exponential model.

8. Show how to use logarithms to transform the data given by $y = ax^p$ to produce a linear relationship.

9. The big idea using logarithms to transform data is that "if a variable grows _____, its _____ grow linearly."

10. Describe how to *achieve linearity* from a power model as explained on page 777.

11. After using a logarithm transformation, what does the *scatter plot* of the data show?